Biases in Farm-Level Yield Risk Analysis due to Data Aggregation

Fehler in der betrieblichen Risikoanalyse durch Datenaggregation

Robert Finger
ETH Zürich, Switzerland; New Affiliation: Wageningen University, Wageningen, The Netherlands

Abstract

We investigate biases in farm-level yield risk analysis caused by data aggregation from the farm-level to regional and national levels using the example of Swiss wheat and barley yields. The estimated yield variability decreases significantly with increasing level of aggregation, with crop yield variability at the farm-level being up to 2.38 times higher than indicated from national data. Our results show furthermore that inference on shape parameters based on aggregated data might be misleading. Using an example of farm yield insurance, we show that using crop yield variability estimates from aggregated levels leads to erroneous insurance contract specifications.

Key Words

aggregation bias; crop yield variability; risk measurement; data aggregation

Zusammenfassung

Dieser Beitrag untersucht unter Verwendung Schweizer Ertragsdaten für Weizen und Gerste potentielle Fehler, die bei der Analyse von Ertragsrisiken durch die Aggregation von Betriebsdaten auf regionalen oder nationalen Niveaus entstehen. Die geschätzte Ertragsvariabilität sinkt signifikant mit einem steigenden Aggregationsniveau, wobei die Ertragsvariabilität auf Betriebsebene bis zu 2,38-mal grösser seien kann, als dies durch nationale Daten angedeutet wird. Des Weiteren zeigen die Resultate, dass Rückschlüsse über Schiefe und Kurtosis der Erträge, basierend auf aggregierten Daten, irreführend sein können. Mittels eines Versicherungsbeispiels wird zudem gezeigt, dass die Ausgestaltung einer Versicherung, basierend auf aggregierten Daten, zu falschen Ergebnissen führen kann.

Schlüsselwörter

Aggregationsfehler; Ertragsvariabilität; Risikomessung; Datenaggregation

1 Introduction

The measurement, management and modeling of agricultural production risk is one of the most important issues in agricultural economic research. In particular, the level of yield risk faced by a farmer is important for the design of appropriate management and insurance strategies (LOBELL et al., 2007). Because long and consistent series of farm yield records are often not available, farm-level yield characteristics are often approximated with aggregated data because it is usually more readily available and allows for comprehensive statistical analysis (e.g. BIELZA et al., 2008a). However, the use of aggregated data has strong implications for the conclusiveness of the results on farm-level risk estimates. In particular, yield variability at the farm-level is significantly higher than on more aggregated levels (e.g. FREUND, 1956). Thus, it is of particular importance to know what kind (and magnitude) of error is made by using for instance regional or national crop yield data to make inference on crop yield distributions at the farm-level.

An example for aggregation biases (i.e. measurement errors due to the aggregation of farm-level data) using yield observations from 5 wheat producing farms in Switzerland is given in table 1. These 5 farms are taken from one municipality in the canton of Bern. The average of farm-level standard deviations is 8.14 dt/ha[1]. In contrast, the analysis of aggregated data (i.e. taking annual averages over these 5 farms) leads to a smaller standard deviation of 6,17 dt/ha. Yield variability decreases with aggregation because the temporal variations of crops yields at different farms are not perfectly correlated, e.g. due to different soil and management conditions as well as local weather, pest and infrastructure problems (HENNESSY, 2009; LOBELL et al., 2007). Thus, the level of data aggregation and the correlation of crop yields at different

[1] dt denotes decitonnes and ha hectares.

farms (or other units) is essential for the 'degree' of aggregation bias (KNIGHT et al., 2010). In this respect, we assume that aggregation biases increase if the farm-level data in the example presented above would have been taken from different regions, or the data aggregation would have been made over more farms.

Table 2 gives an overview of publications related to the empirical estimation of aggregation biases. It shows that the standard deviation of crop yields on the farm-level can be up to five times higher than on more aggregated (e.g. county, state or national) levels, while the usual aggregation bias is in the range of factor 2. Thus, the use of aggregated data leads to an underestimation of farm-level yield variability. This fact has implications for risk-programming models (e.g. BECHTEL and YOUNG, 1999; DEBRAH and HALL, 1989; ÖNAL and MCCARL, 1989), the up-scaling of crop models (e.g. HANSEN and JONES, 2000) and for the calculation of crop insurance premiums (e.g. COOPER et al., 2009; SCHURLE, 1996; WANG and

Table 1. An example for aggregation biases using wheat yield observations from 5 farms

	1992	1993	1994	1995	1996	1997	1998	1999	2000	Farm-level SD
Farm 1	47.7	43.6	38.7	39.8	62.9	46.2	39.1	40.6	47.0	7.55
Farm 2	49.0	54.2	48.6	62.8	64.4	59.8	60.1	38.9	45.7	8.71
Farm 3	38.8	45.3	43.5	43.7	52.6	57.3	58.4	48.2	42.8	6.83
Farm 4	41.8	41.3	47.9	45.2	59.3	56.3	46.5	36.3	44.6	7.25
Farm 5	44.5	75.4	58.5	62.6	70.7	71.2	59.5	57.8	48.3	10.37
								Average of Farm-Level SD		**8.14**
								SD of Aggregated Yield Data		**6.17**
Average of Farm 1-5	44.36	51.96	47.44	50.82	61.98	58.16	52.72	44.36	45.68	

Yields are given in dt/ha. SD denotes the standard deviation.
Source: Bookkeeping data from the Farm Accountancy Data Network (Agroscope Reckenholz Taenikon Research Station, 2005).

Table 2. Overview on empirical literature related to aggregation biases in crop production

Publication	Considered crops	Region (or country) of the study	Compared data aggregation levels	Estimated aggregation biases (ratio of SD between aggregation levels)
ANTÓN and KIMURA (2009)	barley, oilseeds, rye, sugar beet, wheat	Germany	farm vs. national	2-4.1[a]
BECHTEL and YOUNG (1999)	winter wheat	Washington (United States)	farm vs. county	1.39[a]
CARTER and DEAN (1960)[b]	sugar beets	California (United States)	farm vs. state	2
COBLE et al. (2007)	cotton, maize, soybean	United States	farm vs. county farm vs. state	1.5-1.9[a] 2.1-2.4[a]
COOPER et al. (2009)	maize, soybean, wheat	Illinois, Kansas (United States)	farm-county farm vs. state	1.12- 1.38 1.3-1.9
DEBRAH and HALL (1989)	maize, soybeans, tobacco, wheat	Kentucky (United States)	farm vs. county	1.5-2.8
EISGRUBER and SCHUMAN (1963)[b]	maize, oats, soy-beans, wheat	Indiana (United States)	farm vs. state	1.4-2.0
LOBELL et al. (2007)	wheat (irrigated)	Yaqui Valley, San Luis Rio Colorado Valley (Mexico)	farm vs. regional	1.58
MARRA and SCHURLE (1994)	wheat	Kansas (United States)	farm vs. county	1.3
POPP et al. (2005)	canola, flax, wheat	Manitoba (Canada)	field vs. municipality	1.4 -2.2
RUDSTROM et al. (2002)	wheat	Manitoba (Canada)	field vs. municipality	1.2-5

a) Comparison of relative yield variability (coefficients of variation). SD denotes standard deviation.
b) Data is taken from MARRA and SCHURLE (1994).
Source: own literature review

Table 3. Characteristics of yield distributions for Swiss wheat and barley on the national level for the period 1990-2008

Crop	Characteristics of yield distributions				
	Mean	**SD**	**CV**	**Skewness**	**Kurtosis**
Wheat	58.56	3.53	0.06	0.11	0.53
Barley	59.51	4.74	0.08	0.08	2.41

Crop yields are given in dt/ha. SD and CV denote standard deviation and coefficient of variation (calculated as SD/Mean).
Source: SBV (2010)

ZHANG, 2002). The presence of these aggregation biases might imply that the assessment of risks faced by farmers has to be based on farm-level data, and the use of aggregated data is not appropriate (COBLE and BARNETT, 2008). However, the use of conversion factors might be useful to calculate farm-level risks based on aggregated data (e.g. COOPER et al., 2009).

Table 2 shows that the geographical focus of most studies that empirically address aggregation biases is the North-American continent, mainly motivated by the relevance of agricultural insurances (e.g. COOPER et al., 2009; RUDSTROM et al., 2002). In contrast to North-America, the penetration of insurance in European agriculture is currently low (see e.g. BIELZA et al., 2008a, b, 2009 for an overview). However, increasing risks due to market liberalization (e.g. by increasing price volatility) and climate change (e.g. by increasing yield volatility) might lead to an emphasis of agricultural insurance solutions in Europe (e.g. BIELZA et al., 2009; MAHUL, 2003; MUSSHOFF et al., 2009; TORRIANI et al., 2008). Due to different structures of agricultural production systems in some parts of Europe and in North-America[2], existing results listed above (i.e. the estimated aggregation biases) might not be generally applicable. Furthermore, though higher moments (i.e. skewness and kurtosis) of crop yield distributions become more relevant in agricultural risk analysis (e.g. GROOM et al., 2008), aggregation biases in these shape parameters of crop yield distributions have received little attention so far.

Against this background, the aim of this paper is to empirically address problems associated with aggregation biases in the first four moments of crop yield distributions using the example of Swiss wheat and barley yields. Moreover, we investigate the aggregation effect at the farm-level, i.e. the effect of farm crop acreage on individual production risks. Finally, the implications of aggregation biases at different scales are discussed using an example of farm-yield insurance.

2 Data

Our analysis focuses on time series of wheat and barley, which have been chosen because they are the two most important cereals in Switzerland. National level wheat and barley yields for the period 1990-2008 are taken from the database of the Swiss Farmers Union (SBV, 2010)[3]. Because Swiss cereal yields on the national and regional level do not exhibit any trend since the early 1990s (see FINGER, 2010a, for a discussion), no detrending of crop yield data has been applied. Table 3 presents the mean, standard deviation (SD), coefficient of variation (CV), skewness and kurtosis of wheat and barley yield distributions at the national level. We find no significant differences between wheat and barley mean yields, and the hypothesis of normal distribution is not rejected for both wheat and barley distributions (by the Shapiro-Wilk test). However, we find a significantly higher yield dispersion for barley than for wheat yields.[4] Thus, while expected yield levels of wheat and barley seem to be identical, higher production risks are indicated for barley.

Farm-level bookkeeping data is taken from the Farm Accountancy Data Network (FADN) and covers the period 1990-2008 (AGROSCOPE RECKENHOLZ TAENIKON RESEARCH STATION, 2005). The FADN system of Switzerland consists of approximately 3 500 farms (LIPS, 2009), but the representativeness of

[2] For instance, the average size of farm-level arable land in our Swiss sample is about 10 ha, while MARRA and SCHURLE (1994) analyze farms in Kansas with an average size of about 160 ha.

[3] In addition, the database of the Food and Agriculture Organization (FAO, 2010) was considered to derive crop yields on the national level. The choice of the database did not affect the qualitative interpretation of the results presented in this paper.

[4] Using the Ansari-Bradley test (see e.g. FINGER and STEPHAN, 2010, for details).

the Swiss FADN data is limited due to the sampling methods applied (MEIER, 2005). Data is taken from 12 regions at the Swiss Plateau (LEHMANN, 2010), the main crop production area in Switzerland. These regions are constructed according to the following criteria: a) there has to be crop production within the region, b) a sufficient amount of farms have to be available within the FADN database for a region in order to construct average values at regional levels, c) sufficient detailed weather data (taken from the Federal Office of Meteorology and Climatology, MeteoSwiss) has to be available for this region. Finally, this process envisaged selecting homogeneous regions with respect to climate and production conditions within a specific region, but to ensure heterogeneous conditions between regions. The selected regions are listed in table 4 and reflect the heterogeneity of the Swiss Plateau region with regard to agricultural production systems as well as with regard to climatic conditions.

In our sample, there are 2 723 and 2 851 farms that have records (i.e. at least one entry) for wheat and barley, respectively, for the period from 1990-2008. Based on this full sample, we construct annual crop yields at the regional level. Thus, for each year and region, the average of all available farm-level crop yields is calculated. The resulting 19 years of observations for each region are used in subsequent steps to estimate characteristics of yield distributions at the regional level.

In this full sample, most farms have only a small number of observations in the analyzed period. Particularly, 475, from the total set of 2 723 wheat producing farms, have reported wheat production only for single year. In contrast, only 9 farm-records are complete for the period 1990-2008. A similar picture is found for barley: 458 farms have only a single entry and 14 farms have continuous records. The lack of continuous farm-level data is caused by crop rotation requirements, abandonment, merger or start-up of farms, and finally due to sample selection. A possible strategy to cope with the lack of available data is the reduction of the time-horizon under consideration: focusing on the period 2000-2008 would lead to 48 complete records, while for the period 2004-2008 even 222 complete records are reported for wheat. However, a shortening of the analyzed time period would remove the information on yield variability over the last two decades.

To balance between desired long time series and the lack of continuous records, we selected farms that report data for the specific crop under consideration in at least 10 years within the 1990-2008 period. This procedure resulted in 617 and 606 farms for wheat and barley. Table 5 shows farm characteristics in the full and the reduced sample. In this table, average values over time period 1990-2008 are reported with respect to the area under the specific crop, the total arable land, the farm's elevation as well as the number of observations within the 19 year record per farm. For barley, small (but significant) differences are found in all variables. For wheat, significant differences between the total set and the reduced set of farms are only found for the area under wheat. In general, the farms in the reduced sets tend to have a slightly larger area under wheat and barley, respectively. However, as these observed differences are in the magnitude of less than 5%, we conclude that no

Table 4. Characteristics of the 12 used regions in Switzerland

Region Number	Region Name	Area in km^2	Average rainfall sum February-July (1990-2008) in mm	Standard deviation of rainfall in mm
1	Lake Constance	559	534.22	100.18
2	Schaffhausen	510	475.03	111.58
3	Centre Thurgau	653	656.06	125.87
4	Aargau North	359	537.29	133.81
5	Aargau West	445	550.60	117.71
6	Three Lakes	841	613.47	107.54
7	Bern East	450	590.70	133.44
8	Seeland	308	557.99	158.04
9	Bern Mittelland	503	552.02	131.72
10	Lake Biel	194	505.81	129.45
11	Lake Neuchâtel	567	490.40	119.68
12	Centre Vaud	1 108	543.29	144.65

Regional rainfall sums are constructed as average of all available rainfall measurement stations in the specific region. See LEHMANN (2010) for details and a map of the regions.
Source: LEHMANN (2010), MeteoSwiss, SwissTopo

Table 5. Farm characteristics of the full and the reduced set of farms included in the datasets for wheat and barley

Variable	Mean	Median	Min	Max	SD
	Wheat – Full sample of 2723 farms				
Area under wheat	3.61	2.8	0.52	28.48	2.90
Arable land	10.81	9.14	0.7	64.28	7.35
Elevation	534.70	520.00	310.00	935.00	98.41
Number of available observations[a]	5.93	5	1	19	4.54
	Wheat – Reduced sample of 617 farms				
Area under wheat	3.64***	2.26	0.69	18.33	2.26
Arable land	11.04	9.91	2.13	50.14	5.99
Elevation	520.65	500	320	835	93.26
Number of available observations	12.95	13	10	19	2.40
	Barley - Full sample of 2851 farms				
Area under barley	1.86	1.53	0.51	15.64	1.25
Arable land	9.94	8.28	0.70	64.28	7.18
Elevation	544.80	530.00	310.00	940.00	100.45
Number of available observations	5.85	4	1	19	4.53
	Barley - Reduced sample of 606 farms				
Area under barley	1.93***	1.68	0.61	9.28	0.98
Arable land	9.68***	8.64	1.29	48.93	5.78
Elevation	548.06***	530.00	320.00	835.00	100.55
Number of available observations	13.15	13	10	19	2.52

Farm characteristics are the farm-level averages for the period 1990-2008. Area under wheat and arable land are given in hectare, elevation is given in meters above sea level. SD, Min and Max denote the standard deviation, Minimum and Maximum, respectively. a) Number of available observations per farm (i.e. number of years, out of 19 possible, where the farm has reported data). *** denotes significant differences at the 1% level of significance between the total and the reduced set of farms indicated by the Mann-Whitney-Wilcoxon test.
Source: own calculations

systematic sample selection biases are introduced due to the farm selection process[5].

3 Methodology

To estimate biases in yield distribution characteristics due to data aggregation, we analyze location and scale as well as the shape parameters of crop yield distributions at different aggregation levels. To this end, the mean, standard deviation, skewness and kurtosis of wheat and barley yields are estimated using a 19 year period of observations for every farm and region as well as for the national level. These estimates are compared with each other by constructing ratios between the different levels of aggregation, i.e. estimates for aggregation biases. For instance, for each of the 617 analyzed wheat producers, the yield standard deviation is estimated and compared with the wheat yield standard deviation at the national level by constructing a ratio between these two estimates. The mean of the above described ratios over the 617 farms is reported in table 7 and is referred to as *aggregation bias* throughout this paper.

To test if differences between the different levels of aggregation are significant, we employ tests that have been suggested by WILCOX (2005) using the 'WRS' package from the statistical language and environment R (R DEVELOPMENT CORE TEAM, 2010): In order to compare parameter estimates at the farm or regional level (that include estimates for several farms or regions) with the respective (point) estimate at the national level, we construct bootstrapped confidence intervals of the median of differences. For instance, mean barley yield levels for each of the 606 (barley) farms are compared with the national average yield by calculating the difference of these two values. If data aggregation does not introduce a bias in the estimation of mean yield levels, the median of these differences over all farms is expected to be zero. The bootstrap inference is based on 999 bootstrap samples, which are generated by sampling with replacement from the 606 farm-level estimates. Thus, each of the 999 bootstrap samples leads to a different median of differences between farm-level and national-level mean

[5] Moreover, slightly more observations are available for the first half of the time period analyzed (1990-1999) for both the full and the reduced sample, which is not expected to alter the overall conclusions.

barley yields. The distribution of these 999 different values is employed to construct confidence intervals for the initial estimate, i.e. it is used to assess if the mean barley yield at the farm-level is identical to the national-level mean barley yield.

Comparisons of the farm with the regional level use bootstrapped confidence intervals based on the Yuen-Welch test (YUEN, 1974), i.e. testing for the equality of trimmed means.[6] For instance, the two samples of (606) farm-level and (12) regional-level estimates of average barley yields are compared with each other based on their trimmed means. Bootstrap inference is based on 999 bootstrap samples where both farm- and regional level samples are again generated by sampling with replacement. The here used test procedures (using bootstrap techniques and medians as well as trimmed means) have been selected because they do not rely on distributional assumptions and they are robust against exceptional crop yield observations (i.e. outliers) within the samples.

It has been observed that regional differences in yield variability within a country might be more important than aggregation biases, in particular if these regions face extremely different climatic conditions (GÓRSKI and GÓRSKA, 2003). To analyze the spatial distribution of risks in crop production, we test if the regions are heterogeneous with respect to yield levels and yield variability using the non-parametric (rank-based) Kruskal-Wallis test (to test if expected yield levels differ between the regions, see KRUSKAL and WALLIS, 1952, for details) and the non-parametric Fligner-Killeen test (to test if yield dispersion differs between the regions, see FLIGNER and KILLEEN, 1976, for details), respectively.[7]

An additional analysis conducted in this paper focuses on the determinants of farm-level yield variability. In order to account for different yield levels across farms, coefficients of variation are used. With respect to the determinants of farm-level yield variability, we expect that in particular the area under the specific crop is an important determinant of farm-level crop yield variability (EISGRUBER and SCHUHMANN, 1963; MARRA and SCHURLE, 1994). Thus, a larger area under a specific crop implies a risk reducing effect because yield losses in one field might be smoothed by above average yields on other fields of the farm. Thus, the larger the crop area of a single farm, the smaller should be the negative and positive amplitudes of farm-level yield observations (i.e. reduce yield variability). To estimate the effect of crop acreage on yield variability, we also account for regional differences in crop yield variability, precipitation levels and farm-level pesticide expenditures using a (linear) regression model following MARRA and SCHURLE (1994). Moreover, the number of available observations per farm (out of 19 possible years) is included in the regression model. For wheat and barley, respectively, the regression model is defined as follows:

$$CV_i = \beta_0 + \beta_1 CV_R + \beta_2 RAIN_R + \beta_4 Acreage_i + \beta_5 Pesticide_i + \beta_6 OBS_i + \varepsilon_i$$

Where CV_i, CV_R and $RAIN_R$ denote the coefficients of variation at the farm- and the regional (i.e. the region R where farm i is located) level, as well as the average rainfall sum at the regional level (cp. table 4). Regional rainfall was measured within the growing season (February-July).[8] $Acreage_i$ denotes the average crop-acreage of farm i (e.g. acreage under wheat) and $Pesticide_i$ denotes the average pesticide expenditures of farm i for this crop. Note that all values at the regional level (e.g. the average growing season rainfall sum) are constructed based on 19 years of observations. Farm-level variables (e.g. the average pesticide expenditures) are based on all available observations of farm i within the period 1990-2008. The variable OBS_i denotes the number of available observations per farm (i.e. number of years, out of 19 possible, where the farm has reported data) and ranges from 10 to 19. This variable has been included in the regression to test if the amount of available observations influences yield variability estimates. Finally, ε denotes the disturbance term and β_0 is the regression intercept. Regression analyses are based on 617 and 606 farms for wheat and barley, respectively.

[6] See CONOVER et al. (1981) and WILCOX (2005) for further details. In order to validate the here employed tests, we additionally employed bootstrapped confidence intervals based on other robust measures of location and used non-parametric tests for comparisons (e.g. the Mann-Whitney-Wilcoxon and the Ansari-Bradley test, see FINGER and STEPHAN, 2010, for details). All results remain as the here presented.

[7] We also use the Bartlett Test of Homogeneity of Variances to confirm the results of the Fligner Killeen test.

[8] See table 4 for references. We also considered shorter periods that might be more sensitive to rainfall (e.g. March-June) as well as single months. However, rainfall sums in these periods are highly correlated with each other and led to similar results as the here presented. The variability of rainfall was not considered because it showed high correlation with rainfall sums.

4 Results

4.1 Characteristics of Yield Distributions at Different Aggregation Levels

Table 6 presents characteristics of wheat and barley yield distributions on the farm- as well as regional level. Note that characteristics of yield distributions on the national level are presented in table 3. In these tables, different estimates for the mean, standard deviation (SD), coefficient of variation (CV), skewness and kurtosis of wheat and barley yields are presented. Table 7 presents the ratios of these estimates between different levels of aggregation, and shows test results for differences between the aggregation levels. An example for the interpretation of table 6 and 7 is given in the following: the mean (over all farms) farm-level standard deviation of wheat is 8.30 dt/ha, while the minimum observed farm-level standard deviation is 3.86 dt/ha. In contrast, the average standard deviation of wheat yields at the regional level is 4.65 dt/ha (table 6). The ratio of these estimates for farm- and regional level yield variability is 1.78 (table 7). Thus, wheat yield standard deviation observed at the farm-level is in average 1.78 times higher than the one observed at the regional level.

Table 6. Characteristics of yield distributions for wheat and barley on the farm- and the regional level

	Mean	Median	Min	Max	SD
Wheat					
Farm level[1]					
Yield	57.92	57.89	35.61	82.36	7.23
SD	8.30	7.92	3.86	24.99	2.50
CV	0.14	0.14	0.06	0.45	0.05
Skewness	-0.03	-0.04	-2.47	2.45	0.64
Kurtosis	2.74	2.48	1.31	8.73	1.00
Regional level[2]					
Yield	57.76	58.88	50.86	60.93	3.07
SD	4.65	4.52	3.97	5.56	0.46
CV	0.08	0.08	0.07	0.1	0.01
Skewness	-0.29	-0.45	-0.84	0.49	0.51
Kurtosis	3.29	3.38	2.19	3.94	0.55
Barley	**Mean**	**Median**	**Min**	**Max**	**SD**
Farm level[3]					
Yield	59.87	58.96	34.35	89.48	9.11
SD	11.31	11.01	4.54	21.34	3.01
CV	0.19	0.18	0.08	0.47	0.06
Skewness	-0.12	-0.11	-2.23	2.41	0.60
Kurtosis	2.66	2.38	1.39	9.31	0.95
Regional level[4]					
Yield	61.14	62.12	52.75	65.99	4.21
SD	6.25	6.20	4.80	7.49	0.91
CV	0.10	0.10	0.09	0.12	0.01
Skewness	-0.51	-0.63	-1.15	0.36	0.40
Kurtosis	3.44	3.48	2.43	4.43	0.70

Crop yields are given in dt/ha. 1) Based on the 617 farms with 10 or more years of wheat records. 2) Based on the full set of 2 723 wheat production observations. 3) Based on the 606 farms with 10 or more years of barley record. 4) Based on the full set of 2 851 barley production observations (see section 2 for details). SD, Min and Max denote the standard deviation, Minimum and Maximum, respectively.
Source: own calculations

Table 6 shows that mean yields of wheat and barley are about 58 dt/ha and 60 dt/ha. Mean yields are similar among all aggregation levels, i.e. all ratios presented in table 7 are equal or close to 1. A significant deviation of mean yield levels at the farm- and the national level was indicated only for wheat, though the ratio of these two estimates is 0.99. For all other cases, no significant differences between mean yield levels measured at different levels of aggregation are indicated (table 7). These results show that wheat and barley yield levels that are available at national or regional levels are suitable to predict expected crop yields at the farm-level.

However, the standard deviation and coefficient of variation of wheat and barley yields decreases significantly with increasing levels of aggregation. For wheat, the standard deviation on the farm-level is (in average) 2.35 times higher than on the national level, and 1.78 times higher than on the regional level. These ratios are slightly higher for barley: The standard deviation at the farm-level is 2.38 times higher than on the national level, and is 1.81 times higher than on the regional level.[9] As expected, the aggregation bias increases significantly with the level of aggregation, i.e. from regional to national level. No differences between aggregation biases for absolute yield variability

[9] The minimum farm-level standard deviations of barley yields reported in table 6 show that some farms even have a smaller barley yield variability than it is observed at the national level. However, in average farm-level yield variability is significantly larger than on aggregated levels.

Table 7. Aggregation biases – comparisons of crop yield distributions on farm, regional and national levels

Ratios between:	Mean Yield	SD Yield	CV	Skewness	Kurtosis
Wheat					
Farm/regional	1.00	1.78***	1.79***	0.11	0.83***
Farm/national	0.99**	2.35***	2.32***	-0.27***	5.17***
Regional/national	0.99	1.32***	1.33***	-2.57	6.21***
Barley					
Farm/regional	0.98	1.81***	1.88***	0.24***	0.77**
Farm/national	1.00	2.38***	2.41***	-1.62***	1.10
Regional/national	1.03	1.32***	1.28***	-6.66***	1.43***

Note: *** and ** denote significant differences at the 1% and 5% level. Comparisons with the national level use bootstrapped confidence intervals based on the median, the other comparisons use bootstrapped confidence intervals based on the Yuen-Welch test using 999 bootstrap samples, respectively (see section 3, for details).
Source: own calculations

(i.e. standard deviations) and relative yield variability (i.e. coefficients of variation) are found. In summary, these results suggest stable and clearly directed effects of data aggregation on mean yields (no differences) and crop yield variability (decreasing with the level of aggregation).

In contrast, shape parameters (i.e. skewness and kurtosis) of crop yield distributions indicate no clear direction of aggregation biases, but show significant differences between different levels of aggregation. Farm-level data suggest – in average – no skewness of wheat yields, while negative and positive skewness is indicated at the regional and national level, respectively. Furthermore, our results show leptokurtic distributions on farm- and regional level, but a rather mesokurtic distribution at the national level.[10] These results show that the estimation of shape parameters of crop yield distributions is very sensible to the employed data sources and data aggregation.

4.2 Regional Differences in Crop Yield Variability

In order to analyze the importance of spatial heterogeneity in crop yield variability, we analyze differences in crop yield distributions between regions. Table 8 shows mean yields, standard deviations and coefficients of variation for wheat and barley yields in the 12 study regions. Differences in mean yields across regions are significant, but the null hypothesis that the crop yield dispersions in each of the regions are the same could not be rejected. Coefficients of variation of wheat and barley yields are similar across regions because the observed variations in yield levels and yield variability offset each other to a large extent, i.e. regions with higher yield levels usually also have a larger standard deviation of yields.

Based on these results, we investigate if this heterogeneity of crop yields at the regional level is also reflected in yield variability at the farm-level. To this end, we analyze if the crop yield variability at the farm-level differs between regions. Figure 1 shows boxplots of coefficients of variation for wheat and barley yields at the farm-level separated by region. We employed boxplots that account for (the robust estimated) skewness of data (HUBERT and VANDERVIEREN, 2008), available in the 'robustbase' package of R. The boxplot shows that farm-level coefficients of variation for wheat and barley are in the

Table 8. Characteristics of regional crop yield distributions (1990-2008)

	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12
Wheat												
Mean***	60.71	57.30	58.97	54.31	50.86	56.17	55.45	60.14	59.06	58.79	60.42	60.93
SD (n.s.)	5.12	4.94	5.56	4.86	5.05	4.39	3.97	4.38	4.63	4.20	4.30	4.42
CV	0.08	0.09	0.09	0.09	0.10	0.08	0.07	0.07	0.08	0.07	0.07	0.07
Barley												
Mean***	64.44	61.88	60.19	58.22	52.75	56.73	56.95	65.50	63.24	62.35	65.99	65.43
SD (n.s.)	7.34	5.94	7.49	5.33	4.80	6.47	5.19	5.92	6.62	7.19	7.08	5.59
CV	0.11	0.10	0.12	0.09	0.09	0.11	0.09	0.09	0.10	0.12	0.11	0.09

Crop yields are given in dt/ha. R1-R12 denote Regions 1 to 12 (see table 4 for details). ***) Differences in yield levels for wheat and barley are significant at the 0.01 level indicated by the Kruskal-Wallis test. (n.s.) denotes that there are no significant differences in yield dispersions for wheat and barley indicated by the Fligner-Killeen test.
Source: own calculations

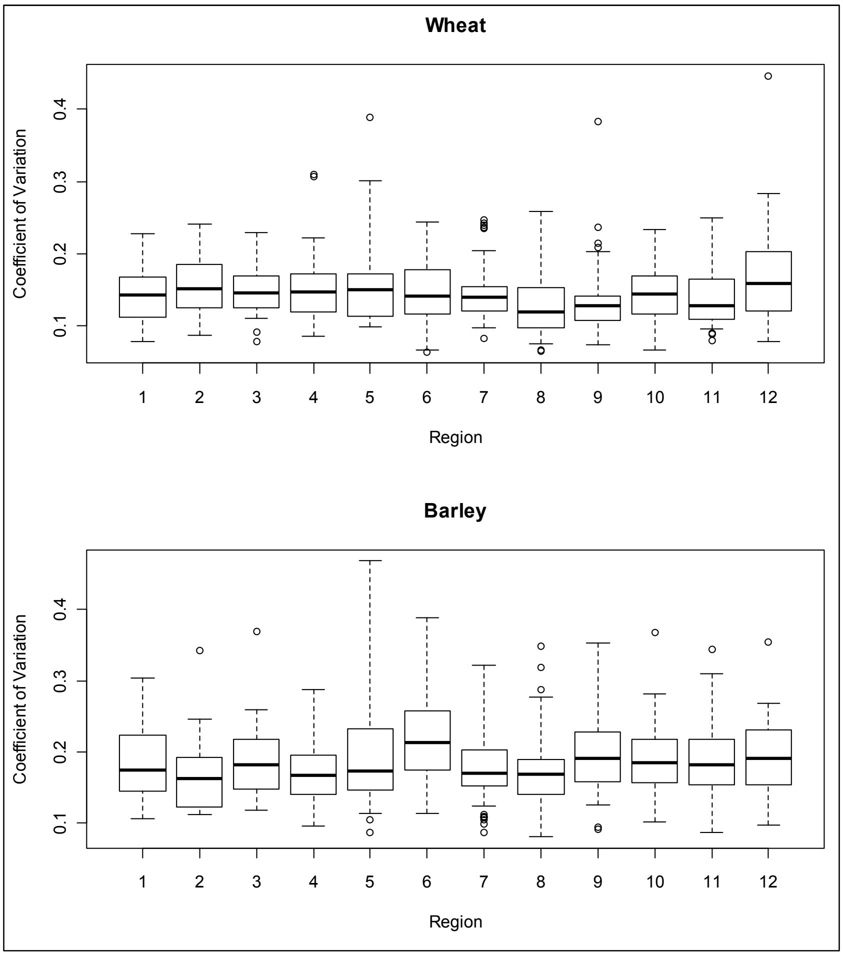
[10] This finding is underlined by different shape parameter estimates at the national level if FAO data is used or a national average is constructed from the bookkeeping data.

same order of magnitude. However, the Kruskal-Wallis test indicates that these differ significantly across regions. This means that there are regional differences in farm-level yield variability.

4.3 Aggregation Biases at the Farm-Level

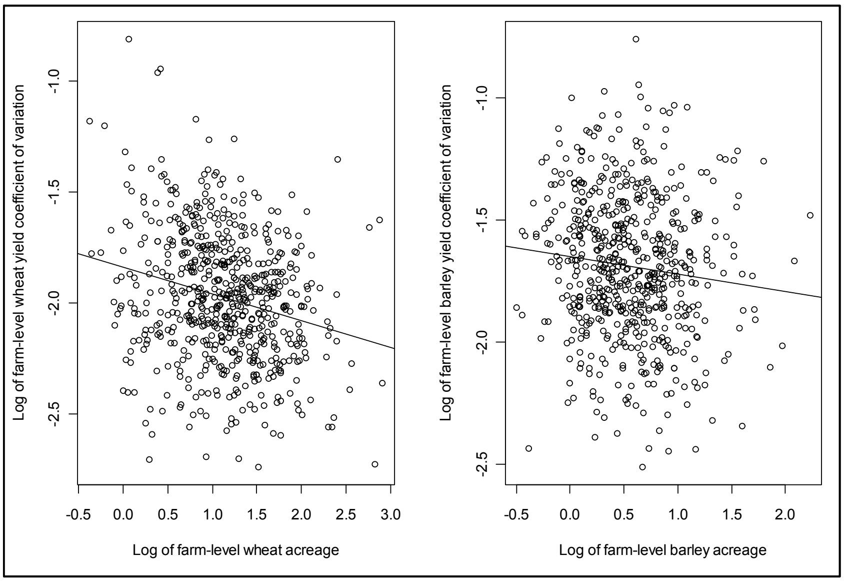
This spatial heterogeneity of farm-level yield variability found above is further investigated using regression analyses. In particular, we aim to test if the area under the specific crop influences the farm-level yield variability (following EISGRUBER and SCHUHMANN, 1963, and MARRA and SCHURLE, 1994). Similar to the aggregation effect if going from the farm to the national level, crop yields at different fields of one farm are not perfectly correlated with each other. Thus, a larger crop production area leads to more smoothed yield variability at the farm-level. Figure 2 shows the relationships of the coefficients of variation of the 617 wheat and 606 barley farms and their respective crop acreage. The data and regression residuals of linear relationships support log-log relationships between crop acreage and yield variability. Thus, yield variability decreases with increasing crop acreage, however, with decreasing marginal effects (cp. MARRA and SCHURLE, 1994). The estimated log-log relationships between coefficients of variation and crop acreage shown in figure 2 can be interpreted as follows: a one percent increase of the area under wheat reduces the farm-level coefficient of variation by 12%. The estimated relationship for barley is, however, smaller. A one percent increase of the area under wheat reduces the farm-level coefficient of variation by about 7%. Note that the range and average of the farm-level area under wheat is larger than for barley (cp. table 5). This

Figure 1. Farm-level coefficients of variation of wheat and barley



Box plots show median (horizontal line), 25th and 75th percentiles (box), whiskers and outliers (circles) are constructed taking the skewness (using the robust skewness estimate medcouple) of the distribution into account (see HUBERT and VANDERVIEREN, 2008, for details).
Source: own calculations

Figure 2. Farm-level coefficients of variation for wheat and barley vs. farm-level crop acreage



Lines show the estimated linear relationship between logarithms of coefficients of variation and crop acreage.
Source: own calculations

smaller barley crop acreage at the farm-level may explain the smaller aggregation effects indicated above. Moreover, this may partially contribute to the generally higher yield variability of barley compared to wheat.

In order to also account for other factors that affect farm-level yield risk (i.e. to 'isolate' the effect of crop acreage), we use a regression approach to analyze farm-level relative yield variability that has been described in section 2. The relationship between farm-level coefficients of variation and explanatory variables is found to be non-linear, and the data and regression residuals of linear relationships support log-log relationships that are employed in the regression approach.[11]

In this regression analysis, regional differences in yield variability (table 8) as well as farm-level pesticide expenditures are considered. Following MARRA and SCHURLE (1994), we also consider the spatial heterogeneity in climatic conditions by including regional average rainfall sums. In addition, we test if the number of available observations per farm (i.e. number of years, out of 19 possible, each farm has reported data) has an influence on the estimated farm-level yield variability.

Table 9 shows that farm-level yield variability differs across regions (only significant for barley). As it has been indicated by figure 1, farm-level yield variability is higher in regions that face generally higher yield risks. Table 9 shows furthermore that increasing rainfall levels lead to increasing yield variability in Swiss barley production. Thus, rather the excess than the shortage of rainfall is the main climatic risk source in Swiss cereal production under current climatic conditions. The crop acreage and pesticide expenditures have the expected risk decreasing effect, though these effects are only significant for wheat. The number of available observations per farm has no influence on farm-level estimates of yield variability. Thus, incomplete time series of crop yield data at the farm-level lead to unbiased estimates of crop yield variability. In summary, these results suggest that increasing crop acreage leads to decreasing crop yield variability at the farm-level. For instance, larger wheat producers usually face lower farm-level wheat yield variability and thus have lower wheat production risks.

Table 9. Log-log regression results: determinants of relative yield variability at the farm-level

	Wheat	Barley
Intercept	-1.308 (-1.30)	-3.736 (-3.67)***
Regional CVs	0.131 (1.22)	0.4540 (4.13)***
Regional rainfall	-0.006 (-0.04)	0.492 (3.23)***
Crop acreage	-0.108 (-5.17)***	-0.041 (-1.53)
Pesticide expenditures	-0.0002 (-2.29)**	-0.0001 (0.50)
Number of available observations	-0.043 (-0.67)	-0.008 (-0.14)
N	617	606
Adjusted R^2	0.06	0.05

Note: *** denotes significance at the 1% level. The dependent variables are farm-level coefficients of variation.
Source: own calculations

[11] Several linear and non-linear specifications have been included in the model selection process following MARRA and SCHURLE (1994).

4.4 Implications of Aggregation Biases: An Example of Yield Insurance

The problem of aggregation biases in insurance applications is illustrated by a farm yield insurance example. Such insurance assumes a "guaranteed" yield level and is used in the USA in the Actual Production History (APH) yield insurance (e.g. KNIGHT et al., 2010). BIELZA et al. (2008a) provide an overview on global applications of farm-yield insurances.

In the here presented example, we assume a coverage level of 90%, i.e. a deductible of 10%. Thus, if the actual crop yield (Y_i) falls below the 90th percentile of average yields ($\bar{Y}$) the farmer is indemnified. This critical yield level Y^C, i.e. below which the farmer gets a payment from the insurance, is defined as follows: $Y^C = 0.9\,\bar{Y}$.

Mean yield levels of 58 dt/ha for wheat and 60 dt/ha for barley are assumed. Because price risks are not considered in this insurance scheme, constant price levels (P_{Crop}) of 53.2 CHF/dt for wheat and 36.7 CHF/dt for barley are used (FAO, 2010).

If the yield Y_i falls below the critical yield level Y^C, we assume that farmers' are indemnified linearly. Thus, the indemnity payment function can be described as follows: $Indemnity = \max\{0, Y^C - Y_i\} \cdot P_{Crop}$.

The premium of this insurance is calculated based on the observed variability of crop yields and results in a 'fair premium', i.e. the insurance premium is equal to the expected indemnity payment (MUSSHOFF et al., 2009). Assuming normally distributed crop yields, we calculate the fair premium by taking standard deviations derived from a) the national, b) the regional, and c) the farm-level. This procedure can reveal the potential influence of data aggregation on insurance premium calculation.

Table 10 shows that the estimated fair insurance premiums differ between the assumed aggregation levels that are used for the estimation of the standard deviations. More specifically, the insurance premiums calculated with national level yield variability are very small. This is due to the fact that standard deviations estimated with the national level crop data indicate only small probabilities of achieving yields below the 90th percentile of average yield levels. In contrast, if crop yield variability is derived from the farm-level data, a 16 times higher fair insurance premium is suggested for wheat. This example shows that the aggregation bias in the estimation of standard deviations has much larger implications if it comes to insurance applications. If farm-level insurance contracts would be specified based on crop yield data from national or regional levels, insurance premiums would not reflect the actual risk at the farm-level. If insurance contracts would be based on aggregated values, expected indemnities at the farm-level would exceed premiums paid and the insurance company would make substantial losses.[12]

Our analysis showed a large heterogeneity of yield variability between farms (table 6). If the farm yield insurances would be specified based on an average farm-level yield risk, only farms with larger than the average risk are expected to buy an insurance contract. Thus, the insurance would be affected by problems of adverse selection, potentially leading to (insurance) market failure because the insurance company is expected to make losses (e.g. BARNETT et al., 2005). In contrast, the use of farm-specific insurance contracts (i.e. based on each farms actual risk) would reduce these problems of adverse selection.[13]

Table 10. Fair insurance premiums (in CHF/ha) based on standard deviations from different aggregation levels

SD based on data at the:	Wheat	Barley
National level	3.98	8.55
Regional level	12.53	20.74
Farm level	63.20	77.94

Source: own calculations

5 Discussion and Conclusion

We investigated biases in farm-level yield risk analysis due to data aggregation from the farm-level to regional and national level using the example of Swiss wheat and barley yields. It shows that the expected (i.e. mean) yields do not differ between different levels of aggregation. However, the estimated variability decreases with increasing level of aggregation, e.g. if comparing the farm- and the national level. Crop yield data at more aggregated levels show lower variability because yields at different farms or regions are not perfectly correlated with each other. Thus, the aggregation of several farms hides the real underlying crop yield variability. Our results show that crop yield variability for wheat and barley at the farm-level is 2.35 and 2.38 times higher than it is indicated from national level data. A smaller aggregation bias for crop yield variability (1.78 and 1.81 times higher standard deviation) has been found between the farm and regional level. The analysis of the shape parameters (skewness and kurtosis) of crop yield distributions suggests significant differences between aggregation levels. For instance, regional crop data suggest negatively skewed distributions, while no skewness is indicated by the farm-level data. Thus, inference on shape parameters of farm-level crop yield distributions based on aggregated data might be misleading. The here observed identicalness of mean yields at different aggregation levels as well as the aggregation biases for standard deviations and coefficients of variation in the range of the factor 2 are generally in line with the results of the studies presented in table 2. In contrast to the results of POPP et al. (2005), we do not find different effects of aggregation on absolute and relative measures of farm-level yield variability.

The potential implications of using crop yield variability estimates from aggregated levels to, for instance, establish insurance contracts at the farm-level have been outlined with an example of farm yield insurance. The probabilities that farm-level yields fall below certain levels are massively underestimated if aggregated data are used. In our example,

[12] This conclusion is based on the assumption that premium calculation and indemnity payment are not based on the same level of aggregation. If this would not be the case (e.g. for an area yield index insurance scheme), no aggregation bias would occur.

[13] Furthermore, we are aware that quality risks can also significantly affect cereal producers, e.g. due to sprouting of cereals (LIPS et al., 2008), which should be considered in future insurance products.

the 2.35 higher standard deviation of wheat yield at the farm than on the national level even leads to about 16 times higher fair insurance premiums. Thus, insurance losses would occur if farm-level insurance contracts would be specified based on crop yield data from national or regional levels.

Based on similar results as found in this paper, COBLE and BARNETT (2008) concluded that the assessment of farm-level risks has to be based on farm-level data only. However, this is usually not possible due to the lack of available (long and complete) time series of crop yield data at farm-level. In contrast, short(er) time series might be more often available but lead to imprecise estimates of yield variability. Inference on such small samples might be additionally very vulnerable to outliers in crop yield data (e.g. FINGER, 2010b). In order to overcome these problems, comprehensive time series of aggregated data (e.g. from the national level) might be used together with a conversion factor that adjusts for different crop yield variability at the farm-level (e.g. COOPER et al., 2009; GOODWIN, 2009; WANG et al., 1998). Because both the mean and the variability of crop yields tend to increase over time (i.e. are non-stationary) for many regions of the world (cp. e.g. HAFNER, 2003), GÓRSKI and GÓRSKA (2003) suggest the use of conversion factors that are based on the coefficient of variation. Thus, the here derived factors of aggregation biases might be applied to adjust insurance contracts in Switzerland.[14]

To alternatively reduce aggregation problems in crop insurance programs, farms might be aggregated according to their properties of crop yield distributions (e.g. RUDSTROM et al., 2002; POPP et al., 2005, and WANG and ZHANG, 2002). In this case, the scope of an insurance product is not determined by administrative or regional boundaries, but farms with similar production conditions and risk properties (i.e. high correlations of crop yields) are grouped together. This avoids aggregation biases and makes this insurance scheme more attractive for farmers (WANG and ZHANG, 2002).[15] However, such "optimal grouping" may lead to high administrative costs. As a compromise, the use of sub-regions (e.g. sub-county levels) might be a promising approach (WANG and ZHANG, 2002).

In our analysis, we also found aggregation effects at the farm-level, showing that an increasing crop acreage (e.g. due to the aggregation of several fields of one farm) leads to decreasing crop yield variability at the farm-level. This effect has also strong implications for farm-level decisions as well as for the scope and design of insurance products. POPP et al. (2005) note that, if such farm-level aggregation bias is important, rather field- level than farm-level yields should be insured. This might reduce incentives of larger farms (with lower risks) to not participate in insurance programs and thus reduce problems of adverse selection. Furthermore, these results indicate that an increase of the crop acreage is a potential risk reducing strategy in crop production. In line with these results, KNIGHT et al. (2010) show that insurance premiums should be adjusted according to the size of the insured unit. For crop insurances in the USA, they propose a discount for an increasing size of the insured unit as well as for an increasing insured level of aggregation (e.g. going from the field to the farm-level). Though farm structure are completely different, the here presented results for Switzerland suggest similar effects as found for the USA (cp. table 2). Thus, the insurance design approach presented by KNIGHT et al. (2010) should be envisaged if new agricultural insurance schemes are established in Switzerland (and potentially in other European countries).

To derive a comprehensive picture of aggregation biases as well as regional heterogeneity in production risks for Swiss crop production at large, additional crops and regions have to be analyzed. Besides their relevance for the here discussed insurance applications, aggregation biases are expected to be relevant in other agricultural applications. For instance, if recommendations of (farm-level) production decisions are based on aggregated data, crop models are calibrated against regional yield data, and risk-programming models are based on aggregated data. In these fields further consideration of biases in farm-level yield risk analysis due to data aggregation is required.

[14] Currently multi-peril insurance in Switzerland (provided by the Swiss hail insurance company) is available for selected risks such as hail, storm, floods etc., but no farm-yield insurance is available yet.

[15] Note that the avoidance of aggregation biases by appropriate grouping also reduces the basis risk for instance for an area yield insurance (WANG and ZHANG, 2002).

References

AGROSCOPE FAT TÄNIKON (2005): Grundlagenbericht 2004. Agroscope Reckenholz Taenikon Research Station, Taenikon, Switzerland.

ANTÓN, J. and S. KIMURA (2009): Farm-level Analysis of Risk, and Risk Management Strategies and Policies: Evidence from German Crop Farms. Paper presented at the

International Association of Agricultural Economists Conference, Beijing, China, August 16-22, 2009.

BARNETT, B.J., J.R. BLACK, Y. HU and J.R. SKEES (2005): Is Area Yield Insurance Competitive with Farm Yield Insurance? In: Journal of Agricultural and Resource Economics 30 (2): 285-301.

BECHTEL, A.I. and D.L. YOUNG (1999): The importance of using farm-level risk estimates in CRP enrollment decisions. Paper presented at the Western Agricultural Association Annual Meeting, Fargo, ND, USA.

BIELZA, M., C.G. CONTO, C. DITTMAN, F.J. GALLEGO and J. STROBLMAIR (2008a): Agricultural Insurance Schemes. Report to the European Commission.

BIELZA, M., C.G. CONTO, R. CATENARO and F.J. GALLEGO (2008b): Regional yield insurance for arable crops in EU-27. Paper presented at the 12th Congress of the European Association of Agricultural Economists, Ghent, Belgium, August 26-28, 2008.

BIELZA DIAZ-CANEJA, M., C.G. CONTE, C. DITTMANN, F.J. GALLEGO PINILLA, and J. STROBLMAIR (2009): Risk Management and Agricultural Insurance Schemes in Europe. EUR – Scientific and Technical Research Series. Ispra, Italy.

CARTER, H.O. and G.W. DEAN (1960): Income, Price, and Yield Variability for Principal California Crops and Cropping Systems. In: Hilgardia 30 (6): 175-218.

COBLE, K.H. and B.J. BARNETT (2008): An assessment of risk exposure in agriculture: A literature review. Organisation for Economic Co-Operation and Development (OECD), Paris.

COBLE, K.H., R. DISMUKES and S. THOMAS (2007): Policy implications of crop yield and revenue variability at differing levels of disaggregation. Paper presented at the annual meeting of the American Agricultural Economics Association, Portland, USA.

CONOVER, W.J., M.E. JOHNSON and M.M. JOHNSON (1981): A comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data. In: Technometrics 23 (4): 351-361.

COOPER, J., M. LANGEMEIER, G. SCHNITKEY and C. ZULAUF (2009): Constructing Farm-level Yield Densities from Aggregated Data: Analysis and Comparison of Approaches. Paper prepared for the Agricultural & Applied Economics Association's 2009 AAEA & ACCI Joint Annual Meeting, Milwaukee, WI, July 26-28, 2009.

DEBRAH, S. and H. HALL (1989): Data Aggregation and Farm Risk Analysis. In: Agricultural Systems 31 (3): 239-245.

EISGRUBER, L.M. and L.S. SCHUMAN (1963): The Usefulness of Aggregated Data in the Analysis of Farm Income Variability and Resource Allocation. In: Journal of Farm Economics 1 (3): 591-597.

FAO (2010): Food and Agriculture Organization of the United Nations (FAO): internet database: http://www.fao.org.

FINGER, R. (2010a): Evidence of slowing yield growth - the example of Swiss cereal yields. In: Food Policy 35 (2): 175-182.

– (2010b): Revisiting the Evaluation of Robust Regression Techniques for Crop Yield Data Detrending. In: American Journal of Agricultural Economics 92 (1): 205-211.

FINGER, R. and A. STEPHAN (2010): Nichtparametrische 2-Stichproben Dispersionstests. In: Wirtschaftswissenschaftliches Studium 39 (5): 254-256.

FLIGNER, M.A. and T.J. KILLEEN (1976): Distribution-Free Two-Sample Tests for Scale. In: Journal of the American Statistical Association 71 (353): 210-213.

FREUND, R.J. (1956): The introduction of risk into a programming model. In: Econometrica 24 (3): 253-263.

GOODWIN, B. (2009): Payment Limitations and Acreage Decisions under Risk Aversion: A Simulation Approach. In: American Journal of Agricultural Economics 91 (1): 19-41.

GÓRSKI, T. and K. GÓRSKA (2003): The effects of scale on crop yield variability. In: Agricultural Systems 78 (3): 425-434.

GROOM, B., P. KOUNDOURI, C. NAUGES and A. THOMAS (2008): The story of the moment: risk averse cypriot farmers respond to drought management. In: Applied Economics 40 (3): 315-326.

HAFNER, S. (2003): Trends in maize, rice, and wheat yields for 188 nations over the past 40 years: a prevalence of linear growth. In: Agriculture, Ecosystems & Environment 97 (1): 275-283.

HANSEN, J.W. and J.W. JONES (2000): Scaling-up crop models for climate variability application. In: Agricultural Systems 65 (1): 43-72.

HENNESSY, D.A. (2009): Crop Yield Skewness and the Normal Distribution. In: Journal of Agricultural and Resource Economics 34 (1): 34-52.

HUBERT, M. and E. VANDERVIEREN (2008): An adjusted boxplot for skewed distributions. In: Computational Statistics & Data Analysis 52 (12): 5186-5201.

KNIGHT, T.O., K.H. COBLE, B.K. GOODWIN, R.M. REJESUS and S. SEO (2010): Developing Variable Unit-Structure Premium Rate Differentials in Crop Insurance. In: American Journal of Agricultural Economics 92 (1): 141-151.

KRUSKAL, W.H. and W.A. WALLIS (1952): Use of ranks in one-criterion variance analysis. In: Journal of the American Statistical Association 47 (20): 583-621.

LEHMANN, N. (2010): Regional Crop Modeling: How future Climate may impact Crop Yields in Switzerland. Master Thesis. ETH Zurich.

LIPS, M. (2009): Full Product Costs on Base of Farm Accountancy Data by Means of Maximum Entropy. Paper presented at the International Association of Agricultural Economists Conference, Beijing, China, August 16-22, 2009.

LIPS, M., R. FINGER and P. CALANCA (2008): Potenzielle Auswirkungen des Klimawandels auf Risiko und Kosten der Weizenernte. In: Yearbook of Socioeconomics in Agriculture 2008: 219-232.

LOBELL, D.B., J.I. ORTIZ-MONASTERIO and W.P. FALCON (2007): Yield uncertainty at the field scale evaluated with multi-year satellite data. In: Agricultural Systems 92 (1-3): 76-90.

MAHUL, O. (2003): Hedging price risk in the presence of crop yield and revenue insurance. In: European Review of Agricultural Economics 30 (2): 217-239.

MARRA, M.C. and B.W. SCHURLE (1994): Kansas Wheat Yield Risk Measures and Aggregation: A Meta-Analysis

Approach. In: Journal of Agricultural and Resource Economics 19 (1): 69-77.
MEIER, B. (2005): Analyse der Repräsentativität im schweizerischen landwirtschaftlichen Buchhaltungsnetz. Messung und Verbesserung der Schätzqualität ökonomischer Kennzahlen in der Zentralen Auswertung von Buchhaltungsdaten. PhD Dissertation No. 15868. ETH Zurich, Switzerland.
MUSSHOFF, O., M. ODENING and W. XU (2009): Management of climate risks in agriculture-will weather derivatives permeate? In: Applied Economics (in press) (2009) DOI: 10.1080/00036840802600210.
ÖNAL, H. and B. MCCARL (1989): Aggregation of heterogeneous firms in mathematical programming models. In: European Review of Agricultural Economics 16 (4): 499-513.
POPP, M., M. RUDSTROM and P. MANNING (2005): Spatial Yield Risk Across Region, Crop and Aggregation Method. In: Canadian Journal of Agricultural Economics 53 (2-3): 103-115.
R DEVELOPMENT CORE TEAM (2010): R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: http://www.R-project.org.
RUDSTROM, M., M. POPP, P. MANNING and E. GBUR (2002): Data Aggregation Issues for Crop Yield Risk Analysis. In: Canadian Journal of Agricultural Economics 50 (2): 185-200.
SBV (Schweizer Bauernverband) (2010): Statistische Erhebungen und Schätzungen über Landwirtschaft und Ernährung. Various issues. Schweizer Bauernverband (SBV, Swiss Farmers' Union), Brugg, Switzerland.
SCHURLE, B. (1996): The Impact of Size on Yield Variability and Crop Insurance Premiums. In: Review of Agricultural Economics 18 (3): 415-422.
TORRIANI, D.S., P. CALANCA, M. BENISTON and J. FUHRER (2008): Hedging with Weather Derivatives to Cope with Climate Variability and Change in Grain Maize Production. In: Agricultural Finance Review 68 (1): 67-81.
WANG, H.H., S.D. HANSON, R.J. MYERS and J.R. BLACK (1998): The Effects of Crop Yield Insurance Designs on Farmer Participation and Welfare. In: American Journal of Agricultural Economics 80 (4): 806-820.
WANG, H.H. and H. ZHANG (2002): Model-Based clustering for cross-sectional time series data. In: Journal of Agricultural, Biological, and Environmental Statistics 7 (1): 107-127.
WILCOX, R. (2005): Introduction to Robust Estimation and Hypothesis Testing. Academic Press, San Diego, CA.
YUEN, K.K. (1974): The two-sample trimmed t for unequal population variances. In: Biometrika 61 (1): 165-170.

Acknowledgements

This work was supported by the Swiss National Science Foundation in the framework of the National Centre of Competence in Research on Climate (NCCR Climate). I would like to thank the Agroscope Reckenholz-Tänikon Research Station for providing the FADN data, the Swiss Federal Office of Meteorology for providing access to the meteorological database, and Nadja El Benni, Niklaus Lehmann and two anonymous reviewers for helpful comments and suggestions on an earlier draft of this paper.

DR. ROBERT FINGER
Wageningen University
Agricultural Economics and Rural Policy Group
Hollandseweg 1, 6706 KN Wageningen, The Netherlands
e-mail: robert.finger@wur.nl